AD659773

RELIABILITY ASSESSMENT IN
THE PRESENCE OF RELIABILITY GROWTH

A. J. Gross and M. Kamins

D D C

RECEIVED

OCT 1 9 1967

C

*The* RAND *Corporation*

SANTA MONICA • CALIFORNIA

Best Available Copy

MEMORANDUM

RM-5346-PR

SEPTEMBER 1967

RELIABILITY ASSESSMENT IN THE PRESENCE OF RELIABILITY GROWTH

A. J. Gross and M. Kamins

DISTRIBUTION STATEMENT
Distribution of this document is unlimited.

$$\mathcal{7\!he}\ RAND\ \mathcal{Corporation}$$

1700 MAIN ST · SANTA MONICA · CALIFORNIA · 90406

PREFACE

This Memorandum is another expression of RAND's long-term interest in and involvement with reliability assessment of Air Force weapon systems, past, present, and future. It provides a methodology for estimating past, current, and near-term future reliability for systems that can be shown to improve in launch and/or in-flight reliability during their development and early operational phases. This methodology should be directly useful to persons responsible for specifying, from actual test results, the appropriate reliability values to be used in targeting and requirements studies. It may also be helpful to those involved in cost-effectiveness evaluations during development and early operational periods, in addition to being of interest to mathematicians, statisticians, operations researchers, and some project or development engineers.

SUMMARY

The relatively brief history of rocket vehicle, and particularly ICBM development, has caused a rediscovery of one of the better-known features of the fly-fix-fly methods of aircraft development--the systems tend to become more reliable as one gains experience and applies it to design improvement. Since changes in reliability have important implications for those involved in planning, procurement, support, and command, a method for assessing this changing reliability at any given stage or projecting it to near-future time periods should be of considerable use.

This Memorandum proposes four reliability growth models or patterns that can be fitted to actual experience data (i.e., launch or flight-test results) to discern the quantitative characteristics of the growth within relatively well-defined tolerances. This objective is achieved by defining appropriate parametric models and subsequently using maximum likelihood procedures to obtain estimates of the parameters, and hence of the reliability. The models are studied in detail with regard to their ability to meet sufficient conditions for the existence of maximum likelihood estimators, and it is shown that only two of them yield maximum likelihood estimates that can be used under the most general circumstances. Numerical procedures are developed for obtaining the estimates of the parameters. Further, the variance-covariance matrix of the estimates is used to construct approximate confidence regions.

These models are compared with each other and with alternative nonparametric and Bayesian approaches, using simulated data to make the comparisons. These comparisons show that under the conditions set forth in this study, three of the parametric models are generally superior in their predictive and assessment characteristics to representative nonparametric methods, and to an applicable Bayesian procedure. However, none of these three parametric models is universally applicable, since the desirable quality of minimum bias can be achieved only by deciding beforehand whether the system reliability is tending closely enough to the usually unattainable goal of 1.0, or perfection, and choosing the model appropriate to the circumstances.

CONTENTS

1. INTRODUCTION

The reliability of a new weapon system is a critical determinant of its effectiveness, and is thus of vital interest to those in the Air Force concerned with strategic planning, procurement, system support, and operational command. The introduction of long-range ballistic missiles into the Air Force weapons arsenal has accelerated the obsolescence of the informal treatment of weapon system reliability characteristics which proved reasonably satisfactory for manned aircraft. The shortcomings were recognized before the missile age, and probably arose from experiences with electronic equipment in the years following World War II. By the time the first generation of ballistic missiles was entering military use, the notion of a mean time between failures (MTBF) had seen wide usage in assessments of alert capability and support requirements, particularly for radars and systems requiring continuous operation. The "one-shot" aspect of ballistic flight has renewed interest in the reliability of independent binomial or Bernoulli trials, but what seems to be a new feature of these systems has added another dimension to the problem.

In May of 1964, Space Technology Laboratories (STL, now known as Systems Division of Thompson-Ramo-Wooldridge) published the results of a study done under contract to NASA, Reliability Growth of U.S. Rockets (U) [9]. The study analyzed the flight test results from nine separate rocket vehicle programs (including four Air Force ballistic missiles), and showed that each one enjoyed a substantial reliability growth during its development and early operational

stages. In an unclassified portion of their report, they concluded
that

> The proportion of successful flights in a program,
> is an indication of the vehicle's average reliability
> for the program. However, the average reliability of
> past flights is not satisfactory for estimating present
> reliability or for predicting the reliability of future
> flights of vehicles.
> If the reliability is increasing from flight to
> flight, the average reliability will lie somewhere be-
> tween the true reliability of the first flight and the
> true reliability of the last flight, where both of the
> true reliabilities are unknown. Present and future
> reliabilities could be grossly underestimated by assuming
> the average reliability of the past.*

STL developed such a prediction model, and applied it to data
from the flight tests of (among others) the Air Force's Minuteman
ICBM, using only the early launches of this system. The projection
of future reliability made from that analysis correctly predicted,
within narrow limits, the outcome of Minuteman flight tests performed
throughout the following two years. Unfortunately, we cannot evaluate,
with regard to subsequent performances, most of the other systems
whose launch results STL analyzed. Many of them had been removed
from operational service by the time the report was published, and
others followed soon after. The data for those that remained opera-
tional are unavailable for this study. Two other Air Force systems
for which we have subsequent flight test data show mixed experience.
For one the STL prediction was excellent, but an equally good one
could have been generated by merely omitting the R & D and early

---

*Reference 9, pages 4-1, 4-2 (U).

operational launches; for the other, subsequent operational results
fell too far from the original prediction to be considered a valida-
tion, but the program was sufficiently unusual in other respects to
make it unsuitable for tests of predictive models anyway.

The purpose of this study is to explore the utilization of
various parametric reliability growth models in evaluating current
reliabilities and predicting near-term future reliabilities of com-
plex weapon systems that exhibit reliability growth during their
development.

The primary emphasis in this study has been on parametric re-
liability growth models rather than on nonparametric ones for the
following reasons:

1. Lower confidence bounds associated with nonparametric
   reliability growth models are very conservative. Since
   important decisions are made on the basis of a lower
   confidence bound, unnecessary penalties imposed upon
   this quantity cause indefensibly higher costs with
   virtually no added capability.

2. Although criticism of parametric methods is generally
   hard to refute on a theoretical basis, parametric
   techniques are often quite useful in practice. Use
   of a nonparametric approach should be contingent upon
   finding a satisfactory substitute for parametric
   methods in hand.

3. There is analytical support for the expectation of an
   exponential reliability growth characteristic in a
   process which is a reasonable facsimile of weapon
   development testing [8]. In addition, there is con-
   siderable empirical material indicating that for such
   processes, reliability is generally increasing, and
   at a progressively decreasing rate [7], [8], [9], [10], [11].

We have not completely ignored other methods of calculating
lower confidence bounds for reliability growth models. For instance,
the Appendix develops a Bayesian procedure for obtaining confidence
intervals for the reliability growth at each stage. The Bayesian

procedure takes into account the "exogenous" information with regard to the system. We later compare the confidence intervals obtained by both the Bayesian approach and another nonparametric method to those obtained by the parametric approach.

The reliability growth models considered for this study assume that the weapon system's reliability during the k-th stage of testing is a function of the ultimate reliability that would be attained if the number of stages is allowed to approach infinity, and one or more parameters modifying the rate of reliability growth. Specifically, we consider first of all reliability growth models of the form

$$(1) \qquad R_k = R_\infty - \alpha F(k),$$

where R_k is the weapon system's reliability at the k-th stage of development, R_∞ is the ultimate system reliability, $\alpha > 0$ is a parameter that quantifies the amount of growth occurring between stages 1 and ∞, and $F(k)$ is a positive decreasing function of k, characterizing that growth. Lloyd and Lipow [6] discuss one member of this class of models and pay specific attention to the case for which $F(k) = 1/k$. The second class of growth models considered is the exponential class, namely,

$$(2) \qquad R_k = 1 - \alpha_1 e^{-\alpha_2 k},$$

where R_k has the same meaning as in (1), α_1 is a parameter indicating the amount of reliability growth, $0 < \alpha_1 \le \alpha_2$, and $\alpha_2 > 0$ is a parameter measuring the rate of reliability growth. This model is developed and utilized in [9].

In a sense, a third class of parametric models can be considered a generalization of (1) and (2) and is represented as

$$(3) \qquad R_k = R_\infty - \alpha^k,$$

where R_k and R_∞ have the same meanings as in (1), and α is a real number that lies in the open interval $(0, 1)$. α indicates the total amount of growth, i.e., the growth from stage 1 to stage ∞.

The fourth and final class of parametric models treated here is given by

$$(4) \qquad R_k = \alpha_1 e^{-\alpha_2/k},$$

where $0 \le \alpha_1 \le 1$ and $\alpha_2 > 0$.

Section II of this Memorandum develops the estimation procedures for reliability growth models, describes the pertinent restrictions, and develops the equations that permit the estimation of current and near-term reliability, as well as the confidence bounds on the esti-mates. Section III displays the behavior of the two models that appear most promising for reliability assessment in the presence of reliability growth,[*] stresses the shortcomings of each, and shows

---

[*]Models 3 and 4 suffer from a number of important shortcomings, among which are (1) computational (convergence) difficulties, (2) a higher variability of prediction, and (3) the necessity to check for the conditions which guarantee a unique maximum (at times, these con-ditions are not met). Although analysis programs were written and exercised for both these models, the results were sufficiently inferior to those for Models 1 and 2 (and the variations thereon) that no detailed examination of those results has been given here. However, the Appendix does present the analytical developments for all four models.

some comparisons between the parametric and other methods. Section IV gives the conclusions of the study.

The mathematically sophisticated reader should have no difficulty in understanding any portion of the discussion that follows. Those whose background does not include training in probability theory and statistics may find certain equations in Sec. II quite difficult to understand, but might be well advised to follow the narrative in any case. Should even this prove too tedious, the reader may prefer to turn directly to the somewhat less mathematically taxing discussion of the models' experimental behavior in Sec. III, and accept the allegations made there concerning the developments in Sec. II.

II. THE MAXIMUM LIKELIHOOD METHOD FOR ESTIMATING PARAMETERS OF RELIABILITY GROWTH MODELS

GENERAL DISCUSSION

The general analysis of reliability growth models for complex weapon systems proceeds in the following manner: A test program is conducted in N stages; at the k-th stage, s_k successes are recorded in n_k trials. When the final or N-th stage is completed, we want to fit to the data a growth curve which then is used to evaluate current reliabilities and predict near-term future reliabilities.

The general parametric reliability growth function can be written as

$$(5) \qquad R_k = f(R_\infty, \alpha_1, \alpha_2, \ldots, \alpha_p; k),$$

where R_k and R_∞ have the same meanings as in (1), and $\alpha_1, \ldots, \alpha_p$ are p-parameters determining the growth of reliability from stage to stage. The vector $(R_\infty, \alpha_1, \ldots, \alpha_p)$ is constrained to lie in a convex region Γ. As a first step in estimating reliability growth, estimates are required of the (p + 1) parameters $R_\infty, \alpha_1, \ldots, \alpha_p$. This is carried out in general by the method of maximum likelihood, whose estimators are used primarily because of their favorable large sample properties.

THE METHOD OF MAXIMUM LIKELIHOOD

At the k-th stage of testing

$$(6) \qquad pr\{x = s_k\} = \binom{n_k}{s_k} R_k^{s_k}(1 - R_k)^{n_k - s_k},$$

where $\text{pr}\{x = s_k\}$ is the probability of exactly s_k successes in n_k trials. Using (5) and assuming that the test stages are statistically independent, the likelihood function whose logarithm is to be maximized is given by

$$\mathcal{L} = \prod_{k=1}^{N} \binom{n_k}{s_k} \left[f(R_\infty, \alpha_1, \ldots, \alpha_p; k) \right]^{s_k} \left[1 - f(R_\infty, \alpha_1, \ldots, \alpha_p; k) \right]^{n_k - s_k}.$$

We then require the set of values $(\hat{R}_\infty, \hat{\alpha}_1, \ldots, \hat{\alpha}_p)$ which lies in the region so that $0 \le f(R_\infty, \alpha_1, \ldots, \alpha_p; k) \le 1$ and which maximizes $\log_e \mathcal{L}$ in this region. To set a unique maximum by the standard technique of partial differentiation of $\log_e \mathcal{L}$ with respect to R_∞, α_1, α_2, $\ldots$, α_p, it is sufficient to demonstrate that $\log_e \mathcal{L}$ is a strictly concave function of these parameters, and that the maximum occurs in the interior of Γ, i.e., the maximum occurs in the region $0 < f(R_\infty, \alpha_1, \ldots, \alpha_p; k) < 1$, for each k.

A sufficient condition that $\log_e \mathcal{L}$ be a strictly concave function is that the matrix

$$\begin{bmatrix} \dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty^2} & \dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty \partial \alpha_1} & \cdots & \dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty \partial \alpha_p} \\[2ex] \dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty \partial \alpha_1} & \dfrac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1^2} & \cdots & \dfrac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1 \partial \alpha_p} \\[2ex] \vdots & \vdots & \cdots & \vdots \\[2ex] \dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty \partial \alpha_p} & \dfrac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1 \partial \alpha_p} & \cdots & \dfrac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_p^2} \end{bmatrix}$$

be negative definite. Thus, assuming that $\log_e \mathcal{L}$ is concave,
$f(R_\infty, \alpha_1, \ldots, \alpha_p; k)$ is a differentiable function of $(R_\infty, \alpha_1, \ldots, \alpha_p)$,
and that the maximum occurs in the interior of Γ, then

$$(7) \quad \log_e \mathcal{L} = \text{const.} + \sum_{k=1}^{N} s_k \log_e f(R_\infty, \alpha_1, \ldots, \alpha_p; k)$$

$$+ \sum_{k=1}^{N} (n_k - s_k) \log_e [1 - f(R_\infty, \alpha_1, \ldots, \alpha_p; k)],$$

$$(8) \quad \frac{\partial \log_e \mathcal{L}}{\partial R_\infty} = \sum_{k=1}^{N} \frac{s_k f_{R_\infty}(R_\infty, \alpha_1, \ldots, \alpha_p; k)}{f(R_\infty, \alpha_1, \ldots, \alpha_p; k)}$$

$$- \sum_{k=1}^{N} \frac{(n_k - s_k) f_{R_\infty}(R_\infty, \alpha_1, \ldots, \alpha_p; k)}{[1 - f(R_\infty, \alpha_1, \ldots, \alpha_p; k)]},$$

$$(9) \quad \frac{\partial \log_e \mathcal{L}}{\partial \alpha_j} = \sum_{k=1}^{N} \frac{s_k f_{\alpha_j}(R_\infty, \alpha_1, \ldots, \alpha_p; k)}{f(R_\infty, \alpha_1, \ldots, \alpha_p; k)}$$

$$- \sum_{k=1}^{N} \frac{(n_k - s_k) f_{\alpha_j}(R_\infty, \alpha_1, \ldots, \alpha_p; k)}{[1 - f(R_\infty, \alpha_1, \ldots, \alpha_p; k)]},$$

where

$$(10) \quad f_{R_\infty}(R_\infty, \alpha_1, \ldots, \alpha_p; k) = \frac{\partial f(R_\infty, \alpha_1, \ldots, \alpha_p; k)}{\partial R_\infty},$$

and

(11) $\quad f_{\alpha_j}(R_\infty, \alpha_1, \ldots, \alpha_p; k) = \dfrac{\partial f(R_\infty, \alpha_1, \ldots, \alpha_p; k)}{\partial \alpha_j}, \quad j = 1, 2, \ldots, p.$

The vector $(R_\infty, \alpha_1, \ldots, \alpha_p)$ for which

(12) $\quad \dfrac{\partial \log_e \mathcal{L}}{\partial R_\infty} = \dfrac{\partial \log_e \mathcal{L}}{\partial \alpha_1} = \ldots = \dfrac{\partial \log_e \mathcal{L}}{\partial \alpha_p} = 0$

yields the maximum likelihood estimators $\hat{R}_\infty, \hat{\alpha}_1, \ldots, \hat{\alpha}_p$ of the parameters $R_\infty, \alpha_1, \ldots, \alpha_p$, respectively.

In general, the system of equations (12) can only be solved by iterative methods so that initial estimates $\hat{R}_{\infty,o}, \hat{\alpha}_{1o}, \ldots, \hat{\alpha}_{po}$ are needed for the iteration scheme. Often the initial estimates are obtained by the "least squares" method. That is, we minimize the sum of squares Ψ of deviations of the observed success ratios s_k/n_k from their expected success ratios $f(R_\infty, \alpha_1, \ldots, \alpha_p; k)$, with respect to the parameter vector $(R_\infty, \alpha_1, \ldots, \alpha_p)$. Thus, we have

(13) $\quad \dfrac{\partial \Psi}{\partial R_\infty} = -2 \sum_{k=1}^{N} [s_k/n_k - f(R_\infty, \alpha_1, \ldots, \alpha_p; k)] f_{R_\infty}(R_\infty, \alpha_1, \ldots, \alpha_p; k),$

(14) $\quad \dfrac{\partial \Psi}{\partial \alpha_i} = -2 \sum_{k=1}^{N} [s_k/n_k - f(R_\infty, \alpha_1, \ldots, \alpha_i; k)] f_{\alpha_j}(R_\infty, \alpha_1, \ldots, \alpha_p; k),$

$i = 1, 2, \ldots, p.$

Among those vector sets for which

$$\frac{\partial \Psi}{\partial R_\infty} = \frac{\partial \Psi}{\partial \alpha_1} = \dots = \frac{\partial \Psi}{\partial \alpha_p} = 0,$$

we find the vector set $(\hat{R}_{\infty,o}, \hat{\alpha}_{1o}, \dots, \hat{\alpha}_{po})$ that is the least squares estimator of $(R_\infty, \alpha_1, \dots, \alpha_p)$.

Iteration schemes are discussed in more detail as we analyze each proposed model in the Appendix. We discuss in particular the two-dimensional Newton s Method used to obtain numerical solutions of the maximum likelihood equations in several of the models.

Let us set $R_\infty = \alpha_o$ for the remainder of this discussion. The maximum likelihood estimators $\hat{\alpha}_o, \hat{\alpha}_1, \dots, \hat{\alpha}_p$ of $\alpha_o, \alpha_1, \dots, \alpha_p$, respectively, are jointly normally distributed when the sample size $m = \sum_{k=1}^{N} n_k$ is large, provided the following conditions are satisfied:

(1) $E_{s_k}\left[\dfrac{\partial \log_e g}{\partial \alpha_i}\right] = 0,$ $i = 0, 1, \dots, p$, where

$$g = \binom{n_k}{s_k}\left[f(\alpha_o, \alpha_1, \dots, \alpha_p; k)\right]^{s_k}$$

$$\left[1 - f(\alpha_o, \alpha_1, \dots, \alpha_p; k)\right]^{n_k - s_k}, \qquad k = 1, 2, \dots, N.$$

(2) $E_{s_k}\left[\dfrac{\partial^2}{\partial \alpha_i \partial \alpha_j} \log_e g\right]^2 < \infty,$

$i, j = 0, 1, \dots, p, \; k = 1, 2, \dots, N.$

(3) $\dfrac{1}{M} \dfrac{\partial^3 \mathcal{L}(\alpha_o, \alpha_1, \ldots, \alpha_p; k)}{\partial\alpha_i \partial\alpha_j \partial\alpha_t}$ is bounded for all possible

values of $\alpha_o, \alpha_1, \ldots, \alpha_p$, $i, j, t = 0, 1, \ldots, p$, and

at each stage k.

(4) $\left. \dfrac{\partial \log_e \mathcal{L}}{\partial\alpha_i} \right|_{\alpha_i = \hat{\alpha}_i} = 0,$ $i = 0, 1, \ldots, p.$ That is to

say, the derivative of $\log_e \mathcal{L}$ vanishes at its maximum

$(\hat{\alpha}_o, \hat{\alpha}_1, \ldots, \hat{\alpha}_p)$.

Assuming that conditions (1)-(4) hold, the maximum likelihood estimators

have the approximate joint normal density with means $\alpha_o, \alpha_1, \ldots, \alpha_p$ and

variance-covariance matrix

$$
\sum = \begin{bmatrix}
\sigma_o^2 & \sigma_{o1} & \cdots & \sigma_{op} \\
\sigma_{o1} & \sigma_1^2 & \cdots & \sigma_{1p} \\
\vdots & \vdots & \cdots & \vdots \\
\sigma_{op} & \sigma_{1p} & \cdots & \sigma_p^2
\end{bmatrix}
$$

where,

$$
\sum{}^{-1} = \begin{bmatrix}
-E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial\alpha_o^2}\right) & -E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial\alpha_o\partial\alpha_1}\right) & \cdots & -E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial\alpha_o\partial\alpha_p}\right) \\
 & -E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial\alpha_1^2}\right) & \cdots & -E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial\alpha_1\partial\alpha_p}\right) \\
 & & & -E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial\alpha_p^2}\right)
\end{bmatrix}
$$

For a further discussion of the large sample properties of maximum likelihood estimators, the reader may refer to Cramér [3, pp. 497-506] and Kendall [5, pp. 1-49].

At this point we note that in some of the simulation studies of Sec. III, condition (5) is violated. This explains the deviation from asymptotic normality which is observed there.

We now derive an approximate 100τ-percent lower confidence limit for R_k--the predicated reliability at the k-th stage of testing. To accomplish this objective we need to obtain an approximate expression for Var $\hat{R}_k$. In the first place, $\hat{R}_k = f(\hat{\alpha}_o, \hat{\alpha}_1, \ldots, \hat{\alpha}_p; k)$, where we assume that $f(\alpha_o, \alpha_1, \ldots, \alpha_p; k)$ is at least twice differentiable in each of the variables $(\alpha_o, \alpha_1, \ldots, \alpha_p)$, and whose second derivatives are bounded for all possible values of $(\alpha_o, \alpha_1, \ldots, \alpha_p)$ at each stage k. We then may approximate Var $\hat{R}_k$ in terms of $\sum$ by expanding $\hat{R}_k$ in a Taylor series about $(\alpha_o, \alpha_1, \ldots, \alpha_p)$ and ignoring terms of order greater than one. Thus,

$$(15) \quad \text{Var } \hat{R}_k \doteq \sum_{i=o}^{p} f_{\alpha_i}^2 \text{ Var } \hat{\alpha}_i + 2 \sum_{o \le i < j \le p} f_{\alpha_i} f_{\alpha_j} \text{ cov } (\hat{\alpha}_i, \hat{\alpha}_j),$$

where

$$f_{\alpha_i}^2 = \left[f_{\alpha_i}(\alpha_o, \alpha_1, \ldots, \alpha_p; k) \right]^2$$

$$f_{\alpha_i} f_{\alpha_j} = \left[f_{\alpha_i}(\alpha_o, \alpha_1, \ldots, \alpha_p; k) \right] \left[f_{\alpha_j}(\alpha_o, \alpha_1, \ldots, \alpha_p; k) \right],$$

and Var $\hat{\alpha}_i$ and cov $(\hat{\alpha}_i, \hat{\alpha}_j)$ are the elements of Σ.

Using the theory developed in this section, we develop in the Appendix the parameter estimators, examine concavity problems, and obtain lower confidence limits for the four models described in this Appendix.

First, by necessity, the decision to allocate a particular trial to one stage of testing or another may be quite arbitrary (the configuration may change very slightly from one trial to the next, obviating any attempt to populate a given stage with homogeneous items). We should thus like to have arbitrary grouping be of little consequence to the resulting estimates. Second, the form of the actual underlying growth curve should have as little effect as practical. In other words, the estimate should not be sensitive to whether a given level of reliability was reached by vigorous early growth, followed by a tapering off to a virtually constant value, or by a slow, sustained growth process. Third, the estimating model should be able to do its job whether the predicted reliability is in the region of 0.5, or near 1.0, or anywhere else, in the reliability spectrum.

Since estimates will be reached through extensive computations, it is desirable that these be reasonably compatible with modern computing methods, that is, with digital computation. Thus, a good model should result in a computing algorithm, or routine, that is unlikely to lead to difficulties such as spurious roots, divisions by zero, logarithms of zero, and other stumbling blocks, or to result in instabilities, or divergences, if iterative methods must be used. Likewise, convergence to the proper answer should be reasonably prompt, with no excessive "hunting."

Finally, the estimating model should bear some strong resemblance to physical reality, and must be compatible with the mathematical interpretation of reliability. For example, the parameters of the model (i.e., the quantities for which we will make numerical estimates)

might be such things as the initial reliability, ultimate reliability,
initial growth rate, etc., and any numerical quantity denoting relia-
bility must take on values neither less than zero, nor more than one.

EVALUATION METHODOLOGY

Few of the foregoing qualities can be implemented by straightforward
analytical efforts, because of the excessive complexities involved. For
this reason we chose to study the behavior and "optimization" of the
reliability growth models through Monte Carlo, or simulation methods.
The most-used procedure was to simulate, on a digital computer, a test
program of 72 trials with a given underlying growth characteristic,
usually exponential. In most cases, the first 12 trial results were
then combined to form the first "stage," the next 12 to form the second
"stage," and so on, giving six groups of 12 trials apiece, with the
results punched on a single IBM card. This process was repeated, either
99 times for rough comparisons, or 999 times for more detailed ones.

These data card decks were then analyzed with different reliability
growth models, and the 99 (or 999) resulting predictions of reliability
in the next (i.e., seventh) stage of testing were used as indicators
of estimating ability. The usual measures of quality were the standard
deviation of the estimates, and the bias, the difference between the
average of the estimates and the "correct" answer from the underlying
growth characteristic used to simulate the data. Another thing considered
was the distribution of the estimates, the Normal or Gaussian being
expected.

The most substantial deviations from the foregoing procedure were
made when studying the effect of grouping, or dividing the 72 trials

into stages. In this case, the identical 72 individual trials were divided into four groups of 18 trials apiece, six groups of 12 trials, and eight groups of 9 trials, as shown in Table 1.

Table 1

EFFECT OF VARIED GROUPING IN 999 PROGRAMS OF 72 TRIALS

Number of Groups	Trials Per Group	Reliability of Stage N+1		Prediction Bias	Standard Deviation
		Mean Estimate	Actual		
4	18	0.771	0.838	-0.067	.0451
6	12	0.793	0.830	-0.037	.0494
8	9	0.807	0.825	-0.018	.0512
72	1	0.808	0.812	-0.004	.0730

The first three lines of Table 1 show: the number of groups and trials per group just mentioned; the actual reliability and the average of 999 predictions for the N+1st stage; the difference (bias) between the latter two, and the standard deviation of the 999 predictions with the generalized model,[*] $R_k = R_\infty - \alpha F(k)$, and with $F(k) = e^{(1-k)/4}$. The fourth line of the table shows a comparison with analysis by the exponential model, $R_k = 1 - \alpha e^{-\beta k}$, where each individual trial constitutes a stage of testing. Because of practical limitations on the function $F(k)$, the generalized model cannot be extended to analyze more than about 12 stages. The solution for the estimates of the parameters of the exponential model suffers convergence difficulties with more than one trial per stage.

---

[*]For purposes of facilitating recognition, Model 1 will henceforth be referred to as the "generalized" model, alluding to the wide range of choices for $F(k)$, while Model 2 will be called the "exponential" model for what should be an obvious reason.

Two major conclusions can be drawn from the table: (1) arbitrary grouping makes relatively little difference in the average prediction, though more numerous groups with fewer trials give slightly higher (and in this case less biased) results; and (2) more numerous groups result in a moderately higher variability of prediction.[*] This raises the question of how large a standard deviation one should reasonably expect under the circumstances. While the answer to that is not easily found, we can get at least some idea from a somewhat different case for which the exact answer is well known.

If, instead of 72 trials with progressively increasing reliability, we had n trials with a constant reliability of p (called Bernoulli or Binomial trials), the standard deviation of $\hat{p}$, the estimate of p, is

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1 - p)}{n}} \, ,$$

which for our case (p = 0.83, n = 72) gives σ = 0.044. Clearly, we are not in as favorable a position as this since the reliability is varying, so we should view with extreme suspicion the values of σ near 0.044, as for example the one for N = 4 groups. Later, we will see further reason for caution in this regard.

Another way of looking at the effect of grouping is to examine a particular series of 72 trials (one of the 999 examined previously), as in Table 2. In this instance (which is not necessarily representative)

Table 2

EFFECT OF VARIED GROUPING ON ONE PROGRAM OF 72 TRIALS

Number of Groups	Reliability Estimates	
	Next Stage R_{N+1}	Ultimate, R_∞
2	0.7122	1.0+
3	0.7480	1.0+
4	0.7820	0.995
6	0.7748	0.8915
8	0.7481	0.8114
9	0.7566	0.8090
12	0.7178	0.7393
72	0.780	--

increased numbers of groups (more than 2 and 3[*]) result in generally lower esimtates, but once again only slightly so. In summary, the simulations show that arbitrary grouping of trials has relatively little effect on the resulting predictions. As we will see a little later, systematic grouping can have somewhat larger and worthwhile effects.

We have just seen an example in which the maximum likelihood estimate given by the generalized model for R_∞, the ultimate reliability, was larger than 1.0, a result that is incompatible with the mathematical restrictions. Fortunately, the likelihood function for this model is always concave downward, so that the so-called "constrained maximum," where R_∞ is required to be less than or equal to 1.0, must occur on the boundary whenever the unconstrained case gives R_∞ above 1.0. Knowing this, we can simply set $R_\infty = 1.0$ in these cases, and recalculate the maximum likelihood estimate $\hat{\alpha}$.

Figure 1 illustrates the three things that happen when this "limitating" process is implemented. The plot shows the ranked values from

---

[*]The exclusion of the two and three-group cases was made because these both resulted in estimates for R_∞, the ultimate reliability, in excess of 1.0, a topic that will be addressed next.

the estimate of R_{N+1}) values from 99 test programs of 72 trials each, where both the underlying growth and the analysis followed the generalized growth model with $F(k) = \frac{2}{1+k}$. (This type of plot, on normal probability paper, is frequently used to show the relationship of test results to the Normal distribution.) The solid dots are the ranked predictions where R_∞ is not restricted, and show the expected normality (the straight line). The open dots show the results where limiting has been implemented, indicating that: (1) the results are no longer normal, thus complicating confidence calculations; (2) the average result has been biased downward by reducing the high estimates while not affecting the low ones; and (3) the standard deviation, a measure of the variability of prediction, has been reduced.

The first of these effects is clearly detrimental, since it counteracts one of the desirable features of maximum likelihood estimation. The second is beneficial in this instance; without limiting, the estimates are biased high. However this is not always the case, as we shall see. The third effect, the reduction in standard deviation, is generally desirable, but we note that the effect occurs entirely because of the reduced estimates at the higher end of the spectrum, thus giving rise to the other two noted features. It should be evident that when the estimate of R_∞ is above 1.0, one might be well-advised to choose a different $F(k)$ or a different model, rather than to follow the procedure mentioned earlier.[*]

---

[*]One should not get the impression that the model will be chosen to fit the data or chosen after the data have been examined. The form of the model is not changed. We change only one of the parameters so the model meets physical constraints, i.e., $0 < R_\infty \leq 1.0$. Later in this section, one such change will be discussed in detail.
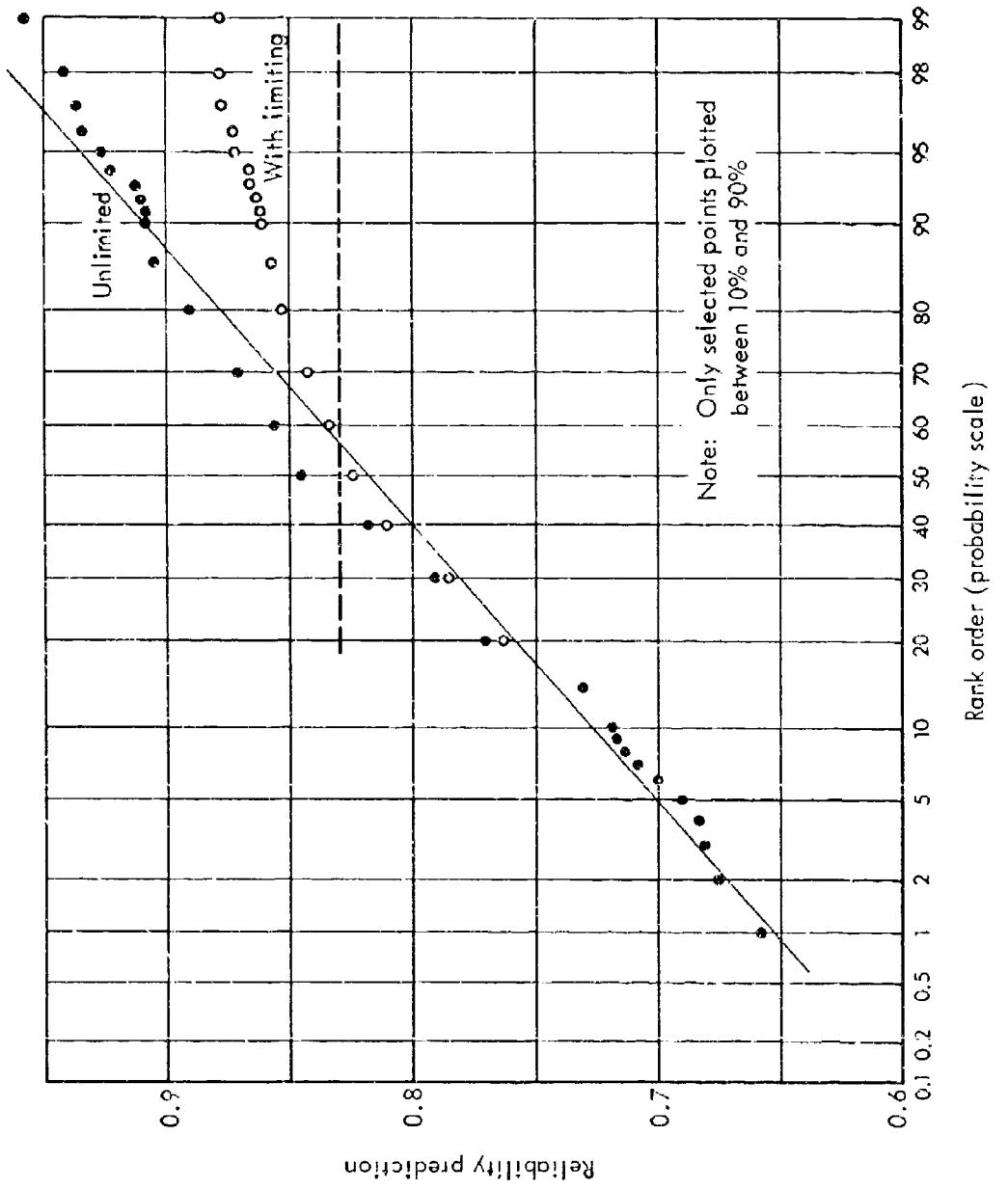
Fig. 1 — Effect of limiting $(R_\infty \leq 1.0)$

Before making some direct comparisons between estimating models, one additional refinement applicable to the generalized model will be discussed. In a reliability research study for NASA, Barlow and Scheuer [1] suggested a method for obtaining maximum likelihood estimates for past or current (but not future) stages of testing in the presence of reliability growth. One feature of their procedure is a regrouping process, whereby in a series of testing stages, adjacent stages are combined wherever the success ratio (successes divided by trials) in the later stage is lower than in the earlier stage. The process is continued until all such "reversals" are eliminated. This process is of substantial benefit to the quality of estimates made with the generalized model.

Figure 2 is a bar chart intended to illustrate both the incidence and the size of the benefits achieved when this process was applied to the data from 99 test programs with underlying hyperbolic reliability growth before using the generalized model. In each case, the data originally consisted of six stages, each having 12 trials. Several of the 99 programs had no reversals, and thus still had six stages after processing. The leftmost bar in Fig. 2 shows that the standard deviation remained unchanged for these programs. The next bar shows that those programs that had one reversal enjoyed a slight reduction in σ, from 0.039 to 0.036 (the shaded area); the remainder of the chart shows how programs with two, three, and four reversals fared. Clearly, before the process was applied, the programs with more reversals had a higher variability of prediction than those with fewer or no reversals, but regrouping according to the Barlow-Scheuer procedure
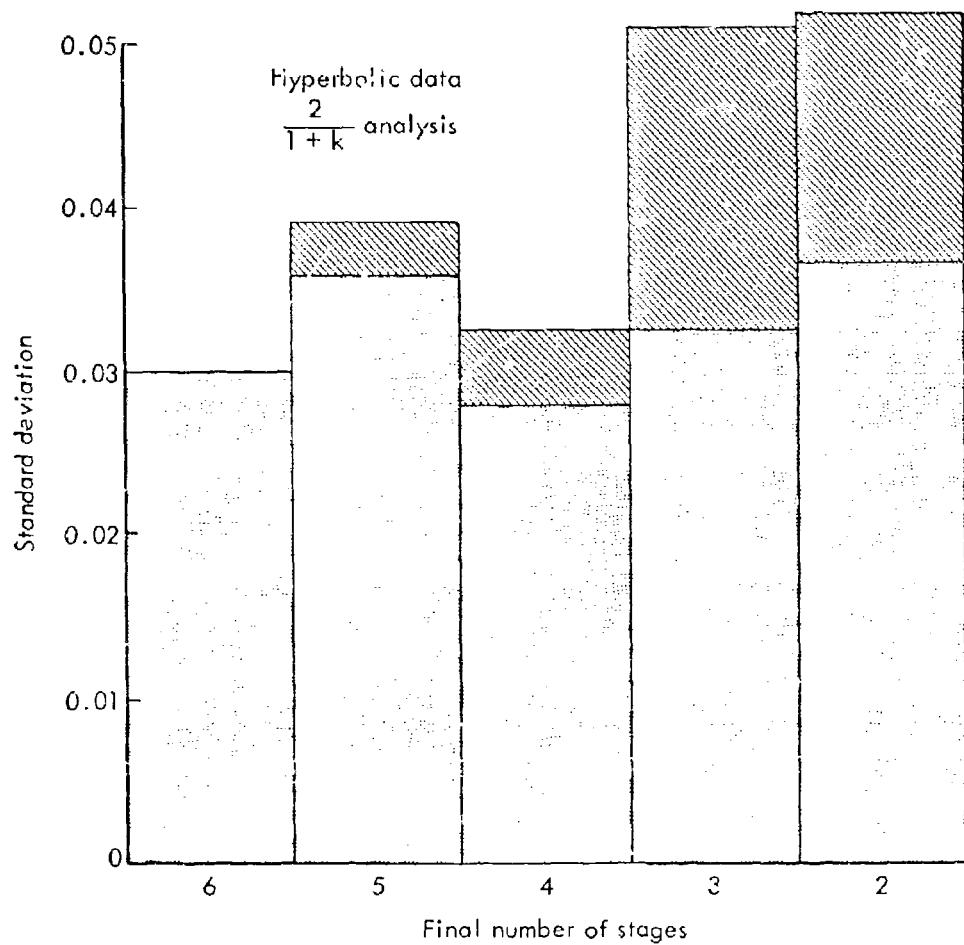
Fig. 2 — Effect of eliminating reversals (per RM-4317-1-NASA)

removed this disadvantage. Although the reasons for expecting benefits

by applying this regrouping process are intuitive in this case, and

based on generally empirical observations concerning the behavior of

maximum likelihood estimators, the fact remains that the process usually

improves estimates where reversals of success ratio are present.

Since the purpose of the Barlow-Scheuer effort is reliability

assessment in the presence of reliability growth, one might reasonably

ask why we did not use their general method rather than just one feature

of it. There are basically four reasons why we chose to take a new

approach.

1. Application of the trinomial model requires assignment of
   failures to "inherent" or "assignable cause" categories,
   something which is often simply impossible.

2. Stages should either be homogeneous or end with an assign-
   able cause failure, both hard to satisfy.

3. There is no way to extrapolate to the N+1st stage.

4. The confidence bound is inadequate, penalized too much
   by early test results.

Of these, only the latter two are critical, since relaxation of the

first two is possible within the framework of the methodology. Further,

the third may be less important late in a test program, provided growth

has substantially abated.

The reader should now have sufficient background to appreciate

the performance differences between the two principal candidate models,

the exponential and the generalized hyperbolic. These models were

used to analyze data generated from three substantially different

growth curves, each having a reliability in the seventh (i.e., next)

stage of 0.8 to 0.85. The first growth curve was exponential with

to the type of data is lower, as indicated by the final line of entries. (The spread of deviations is more than proportionately larger for the exponential.) Finally, the net bias is routinely negative (i.e., low, or conservative estimates) by contrast with the potentially large positive bias of the exponential model when used to analyze vigorous growth data.

Two potential drawbacks of the generalized model are also evident. For exponentially generated data, it is biased where the exponential model is not, and there are valid theoretical reasons (though no demonstrable evidence) to expect exponential data to be more common than other kinds. Also, the standard deviation is once again suspiciously low, though not as much so as the figures would indicate. The 99 program runs used here for analysis had somewhat less variability (in terms of the number of successes in each stage) than would normally be expected. The figures in parentheses are for a 999 program set of data, which were more representative, and which confirm the superiority of the generalized model at a somewhat more realistic level of standard deviation.

A graphical representation offers another means of comparison. Figure 3 shows the 999 predictions from both the exponential and generalized models plotted according to mean rankings on normal probability paper. The lower standard deviation (smaller slope) for the generalized model is quite evident. The bias of the generalized model stands out even more clearly; in approximately nine cases out of ten, the generalized model gives a numerical value below the exponential

-28-
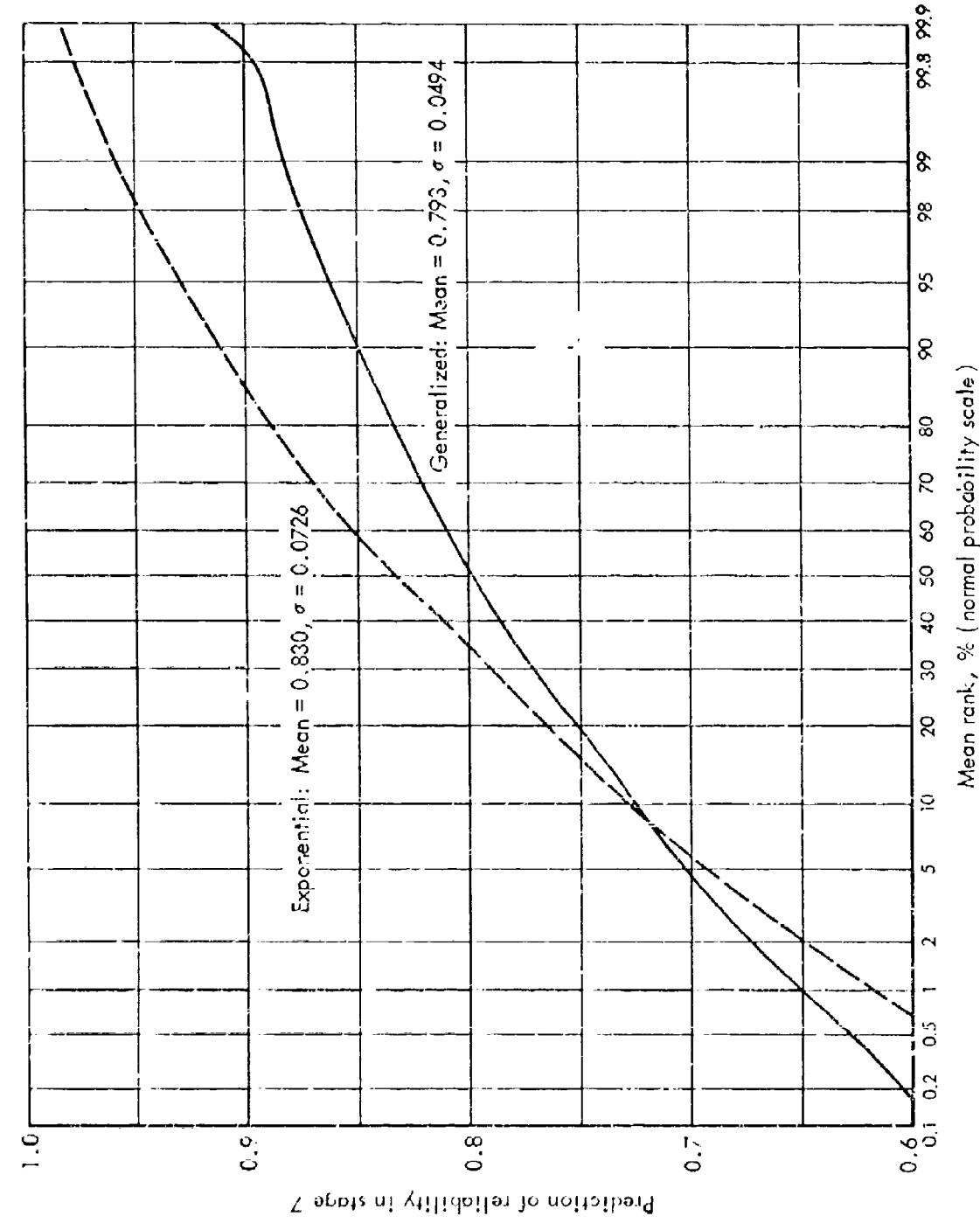
Fig.3 — A comparison of generalized and exponential models for 999 cases

model, and in only one case of the 999 did the value exceed 0 90. The deviation from a normal distribution is also quite evident.[*]

A specific (and quite unrepresentative) example from the 999 program run may serve to dramatize the concern with a "realistic" standard deviation, and to pinpoint the potentially serious shortcoming of this particular version of the generalized model. The solid line in Fig. 4 shows the underlying growth characteristic (reliability versus stage of testing) used to simulate the 999 test programs. The solid dots show what was probably the most unusual of the 999 results, with experience in the first three stages considerably below the expected success ratio, and in the last three stages considerably above the expected. The dashed line shows how the generalized model fits a growth curve to these data, giving a prediction for stage 7 which is quite close to the "correct" answer. The dotted line shows how the exponential model interprets these data, giving a much higher prediction. In spite of the fact that the prediction the exponential model made is substantially further from the true reliability, it should be evident that this higher prediction is eminently more reasonable in view of the experience data itself. In fact, the generalized model with the $e^{(1-k)/4}$ function is simply unable to cope with growth that approaches a reliability of 1.0 at anything more than a slow pace. If a function with somewhat more downward concavity is used (for example, e^{1-k}), the effect is less pronounced.

---

[*]In each case, a beta distribution with the appropriate mean and variance represented the distribution of results far better than the normal. The beta is especially good for the exponential results, and the confidence interval for that model was thus calculated accordingly.
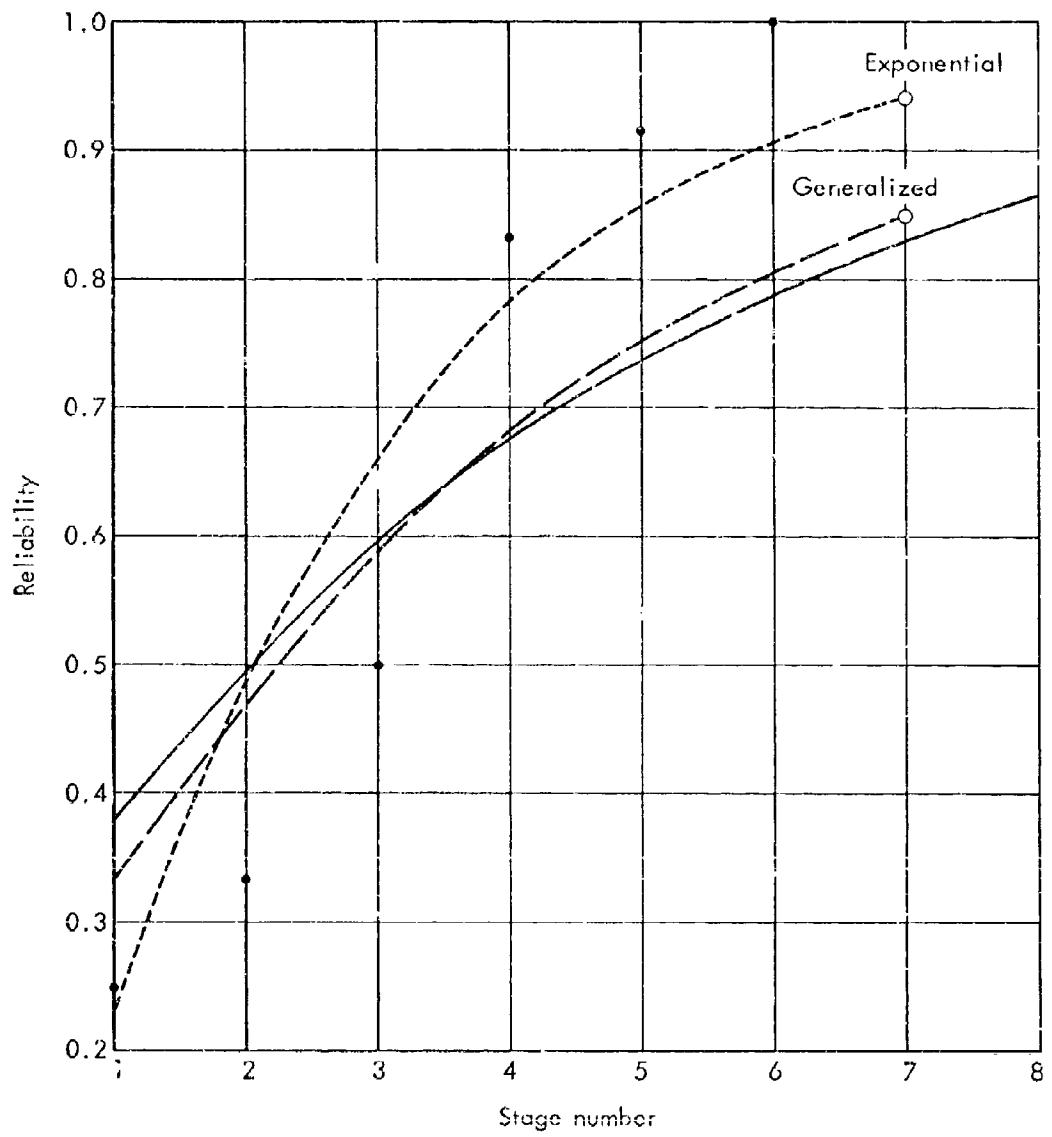
Fig.4 — A particular example from the 999 programs

The shortcomings of the generalized model in analyzing data from a population whose reliability is approaching 1.0 led to the search for another model to deal with this important case. To avoid (at least for the time being) introducing a third estimated parameter, the generalized model was modified in a simple way that would achieve somewhat the same purpose without much extra analytical complication. Briefly, the reliability growth equation is as follows:

$$R_k = R_\infty - \alpha e^{(1-k)/N},$$

where the coefficient N is initially set to an integer between 4 and 8 (usually 6 in our examinations), and both R_∞ and α are estimated. If $\hat{R}_\infty$ turns out to be 1.0 or less, the estimate is accepted. Otherwise, N is reduced by 1.0, and the process is repeated until $\hat{R}_\infty$ drops to 1.0 or below; however, if $\hat{R}_\infty$ remains in excess of 1.0 when N = 1, then the previously described limiting process is introduced, such that

$$R_k = 1 - \alpha e^{1-k}$$

is used to solve for the single parameter $\hat{\alpha}$.

The characteristics of this adaptive model can be appreciated best by comparing them with the exponential model, as in Fig. 5, which shows the ranked predictions for both models when used to analyze the same 999 test programs as before. This time the superiority of the adaptive model over the exponential model is achieved without the drawbacks (negative bias and extreme nonnormality) that characterize
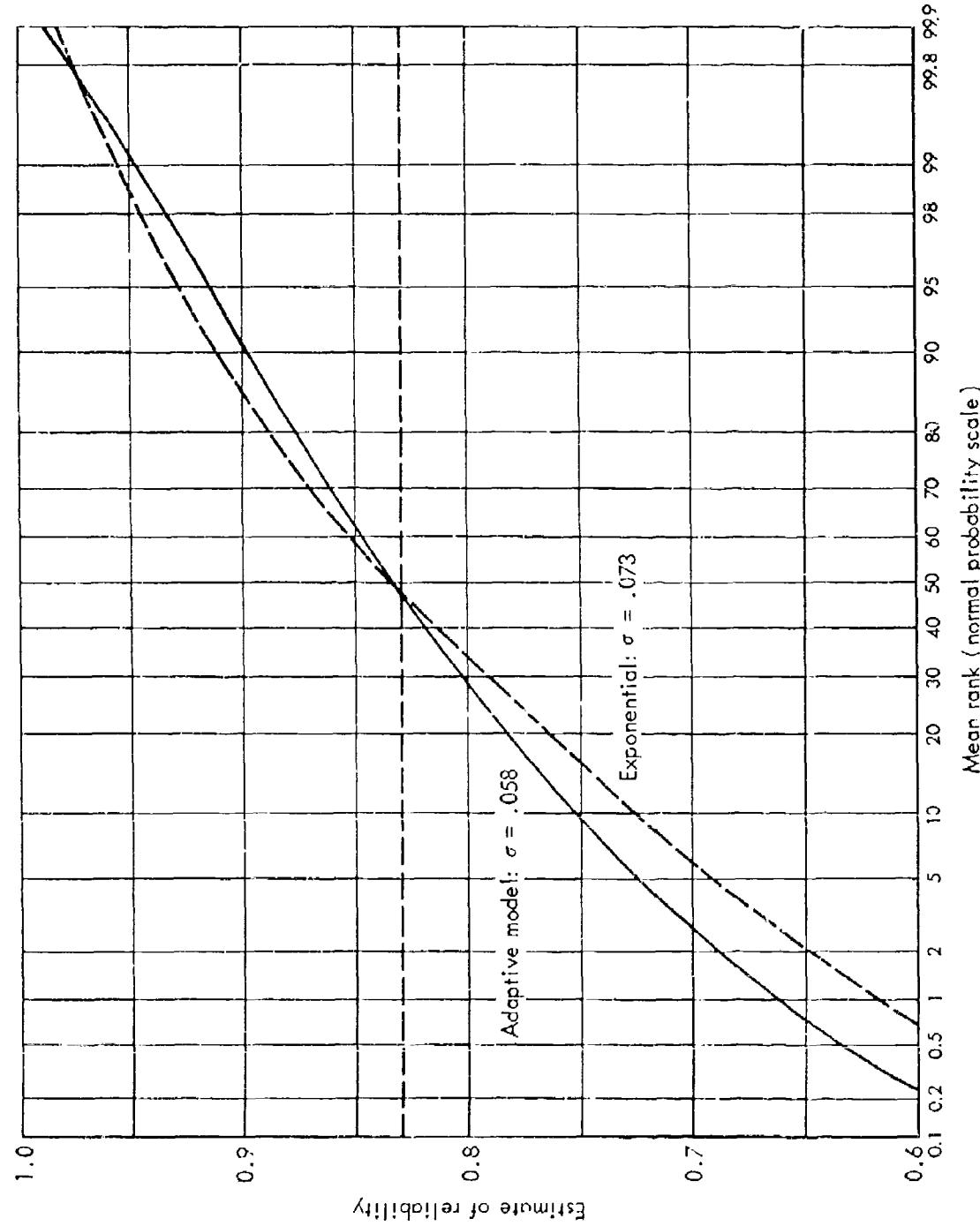
Fig. 5 — Adaptive model: $R_K = R_\infty - \alpha \cdot Exp\left[\left(\frac{1-k}{N}\right)\right]$ (Starting with N = 6, Index N by -1 until $\hat{R}_\infty \leq 1.0$)

the generalized model. This time the result for each model is closely approximated by a beta distribution with matching first and second moments.

In spite of an apparent superiority to the exponential model, the adaptive model does have shortcomings. It shares with the generalized model the inability to make a believable estimate for the case previously described in Fig. 4. (Indeed no other growth function was found that would make an estimate as high as that for the exponential model.) Since that case is anything but typical, it is probably more important to note that the adaptive model also shares a defect of the exponential; it overestimates reliabilities wherever the asymptotic reliability is substantially below 1.0. Thus none of the models described here are universally applicable. One must have some notion regarding the asymptotic reliability if bias is to be avoided (Table 4).

Table 4

EFFECT OF ASYMPTOTIC RELIABILITY ON BIAS FOR THREE MODELS

Asymptotic Reliability	Generalized	Adaptive	Exponential
$R_\infty = 1.0$	Biased low	No bias	No bias
$R_\infty = 0.9$	No bias	Slightly high	Slightly high
$R_\infty = 0.7$	No bias	Very high	Very high

If we are either unwilling or unable to decide whether the asymptotic reliability is near 1.0, then it becomes necessary to introduce a three-parameter model, which estimates the three physical characteristics: ultimate (asymptotic) reliability, starting reliability (alternatively, amount of growth), and a measure of the rate of growth.

-34-

Developing such a model, however, is much more complex than developing the two-parameter models, and will not be attempted here.

SOME ADDITIONAL COMPARISONS

Since nonparametric methods have a strong appeal for use in reliability assessment, our description of the proposed parametric (i.e., assuming an underlying model) methods would be incomplete without a comparison to some pertinent competitors from the nonparametric field. The two of these to be used are the method of Barlow and Scheuer [1], and an extension (described at the end of Sec. II) to a method suggested by Fox [4], involving Bayes' theorem. The latter method was evaluated for the conventional situation where no growth is assumed, and for a second case where a rather accurate representation of the growth between stages was superimposed on the process.

Table 5 shows the results of the calculations applied to the series of 99 cases (with exponential growth), which we previously noted had somewhat less than the expected variance in outcomes. The table shows that for prediction, the adaptive model gives lower variability than any of the nonparametric methods used, and less bias than either of the Bayes approaches. While the Barlow-Scheuer result is actually closer on the average (albeit more variable), it should be noted that this is actually the same result used to assess Stage N (the Barlow-Scheuer method does not include prediction) and is a lucky accident.

In assessing reliability of the most recent stage, the adaptive method shows a clear superiority in both measures. The Barlow-Scheuer results are biased high, largely the result of the substantial number of cases where 11 or 12 successes occurred in the sixth stage, giving

Table 5

A COMPARISON OF EVALUATIONS FOR 99 CASES

Method	Prediction of R(N+1)		Assessment of R(N)		Lower 95% - Bound for R(N)			Conditions
	u	σ	u	σ	μ	σ	x^a	
Adaptive	0.8353	0.0513	0.7898	0.0595	0.6635	0.0690	3	Normal Dist.
					0.6514	0.0667	1	Beta Dist.
Barlow-Scheuer	0.8276	0.0776	0.8276	0.0776	0.529	0.0537	0	
Bayes	0.8399	0.1117	0.7763	0.1036	0.6905	0.0869	11	With Growth
	0.7974	0.1126	0.7363	0.1107	0.6565	0.0948	8	Without Growth
"Correct"	0.83	--	0.789	--	--	--	5	

[a]Number of lower bounds exceeding actual value. Note that while the lower bounds for the adaptive model have no rigid mathematical validity, they still give worthwhile, if conservative, results. This conservatism is typical of both the generalized and exponential models as well, indicating that the estimated variance is substantially larger than the actual in the vast majority of cases examined.

estimates of 0.9167 or 1.0. The Bayes method is biased very low if no growth is assumed, and is still slightly low when a virtually exact representation of the growth is supplied. The adaptive method gives unbiased results, on the average, and with much less variability.

Comparing the lower bounds for the 90-percent confidence interval shows how the adaptive method avoids the excesses of the other two. The Barlow-Scheuer results are excessively conservative. None of the 99 estimates were above the actual, where 5 might normally be anticipated for a valid lower bound. (In a later run of 999 cases, there were still none, where 50 should be expected.) The Bayes results are too optimistic, even though the mean result without growth is almost

identical to the adaptive results. The adaptive method gives results
that are only slightly conservative, and should thus logically be
preferred over the other two.

IV. CONCLUSIONS

On the basis of substantial and independent previous examination
of the general topic (primarily Refs. 7, 8, 9, 10, and 11) the following
conclusions can be drawn:

1. Progressive growth in the reliability of certain types of
   large weapon systems appears to be characteristic of their
   development.

2. Although parametric models require more assumptions con-
   cerning the reliability growth of a system, physical and
   engineering considerations often provide empirical and
   intuitive justification for a characteristic model.

Building on this foundation, the more recent research reported
here permits the following extensions:

3. A simple parametric growth model appears to have advantages
   over other available approaches for assessing reliability.

4. Not all of the parametric models considered are useful, or
   even feasible, because of mathematical difficulties re-
   stricting the choice of parametric form.

5. Parametric growth modeling, in general, permits extrapo-
   lation of previous results to predict near-term future
   reliabilities. In addition, the large sample normality
   properties of maximum likelihood estimation yield a simple
   but effective method of calculating lower confidence
   bounds on reliability for any stage, past, present, or
   future.

6. A given parametric model may yield, in some cases, maximum
   likelihood estimates that lie outside the allowable range
   for determining probabilities. Under these circumstances
   an "adaptive" model has been developed that yields mathe-
   matically as well as physically reasonable results.

7. The models suggested and studied here appear to be rela-
   tively insensitive to several extraneous and usually
   uncontrolled factors such as (1) grouping into stages, and
   (2) the actual form of the underlying growth characteristic.

8. The numerical methods of solution used here are iterative,
   but converge rapidly to the appropriate solution. These
   methods are easily implemented on modern digital computers.

9. Suitable models are available not only for data that yield
   reliabilities approaching 1.0, but also for data that yield
   reliabilities converging to values considerably less than
   unity. The differences appear to be small (by what seem
   to be reasonable standards) but probably deserving
   of attention.

Appendix

ANALYSIS OF RELIABILITY GROWTH MODELS

MODEL 1

This model is covered extensively in Lloyd and Lipow [6] for the case $F(k) = 1/k$. We treat the more general case here in which

$$(16) \qquad f(R_\infty, \alpha, k) = R_\infty - \alpha F(k).$$

The function

$$(17) \qquad \log_e \mathfrak{L} = \sum_{k=1}^{N} \left\{ \log_e \binom{n_k}{s_k} + s_k \log_e [R_\infty - \alpha F(k)] \right.$$
$$\left. + (n_k - s_k) \log_e [1 - R_\infty + \alpha F(k)] \right\}$$

is concave in (y, R_∞), since

$$(18) \qquad \frac{\partial^2 \log_e \mathfrak{L}}{\partial R_\infty^2} = -\sum_{k=1}^{N} \left\{ \frac{s_k}{[R_\infty - \alpha F(k)]^2} + \frac{(n_k - s_k)}{[1 - R_\infty + \alpha F(k)]^2} \right\},$$

$$(19) \qquad \frac{\partial^2 \log_e \mathfrak{L}}{\partial R_\infty \partial \alpha} = \sum_{k=1}^{N} \left\{ \frac{s_k F(k)}{[R_\infty - \alpha F(k)]^2} + \frac{(n_k - s_k) F(k)}{[1 - R_\infty + \alpha F(k)]^2} \right\},$$

and

$$(20) \qquad \frac{\partial^2 \log_e \mathfrak{L}}{\partial \alpha^2} = -\sum_{k=1}^{N} \left\{ \frac{s_k F^2(k)}{[R_\infty - \alpha F(k)]^2} + \frac{(n_k - s_k) F^2(k)}{[1 - R_\infty + \alpha F(k)]^2} \right\};$$

and the matrix

(21)
$$
\begin{bmatrix}
\dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty^2} & \dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty \partial \alpha} \\[2em]
\dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_\omega \partial \alpha} & \dfrac{\partial^2 \log_e \mathcal{L}}{\partial \alpha^2}
\end{bmatrix}
$$

is negative definite for $N > 1$. If $N = 1$, we have but one stage, and reliability growth cannot be assessed.

Thus, the maximum likelihood estimators in the region $0 < \alpha F(k) < R_\omega < 1$ for $k = 1, \ldots, N$ are found by the usual differentiation techniques:

(22)
$$
\frac{\partial \log_e \mathcal{L}}{\partial R_\omega} = \sum_{k=1}^{N} \frac{a_k}{\hat{R}_\omega - \hat{\alpha}F(k)} - \sum_{k=1}^{N} \frac{n_k - a_k}{1 - \hat{R}_\omega + \hat{\alpha}F(k)} = 0,
$$

and

(23)
$$
\frac{\partial \log_e \mathcal{L}}{\partial \alpha} = -\sum_{k=1}^{N} \frac{a_k F(k)}{\hat{R}_\omega - \hat{\alpha}F(k)} + \sum_{k=1}^{N} \frac{(n_k - a_k)F(k)}{1 - \hat{R}_\omega + \hat{\alpha}F(k)} = 0.
$$

Following Lloyd and Lipow's development, first approximations to $\hat{R}_\omega$ and $\hat{\alpha}$, which we label $\hat{R}_{\omega,0}$ and $\hat{\alpha}_0$, are then given as

(24)
$$
\hat{\alpha}_0 = \frac{\dfrac{1}{N}\left[\sum_{k=1}^{N} a_k/F(k) - \dfrac{\left(\sum_{k=1}^{N} 1/F(k)\right)\left(\sum_{k=1}^{N} a_k\right)}{N} \right]}{\dfrac{1}{N}\left(\sum_{k=1}^{N} 1/F(k)\right)\left(\sum_{k=1}^{N} F(k)\right) - N},
$$

rnd

$$(25) \qquad \hat{R}_{\omega,o} = \frac{\frac{1}{\bar{n}}\left[\left(\sum_{n=1}^{N} F(k)/N\right)\left(\sum_{k=1}^{N} s_k/F(k)\right) - \sum_{k=1}^{N} s_k\right]}{\frac{1}{N}\left(\sum_{k=1}^{N} 1/F(k)\right)\left(\sum_{k=1}^{N} F(k)\right) - N},$$

where $\bar{n} = \sum_{k=1}^{N} n_k/N$. Using the initial values, we can then iterate
on these to obtain the approximate maximum likelihood estimators.
The estimating equations are:

$$(26) \qquad \sum_{k=1}^{N} \frac{s_k}{\Delta_{k,u} n_k F(k)} = \hat{R}_{\omega,u+1} \sum_{k=1}^{N} \frac{1}{\Delta_{k,u} F(k)} - \hat{\alpha}_{u+1} \sum_{k=1}^{N} \frac{1}{\Delta_{k,u}},$$

and

$$(27) \qquad \sum_{k=1}^{N} \frac{s_k}{n_k \Delta_{k,u}} = \hat{R}_{\omega,u+1} \sum_{k=1}^{N} \frac{1}{\Delta_{k,u}} - \hat{\alpha}_{u+1} \sum_{k=1}^{N} \frac{F(k)}{\Delta_{k,u}},$$

where

$$(28) \qquad \Delta_{k,u} = \frac{1}{n_k F(k)} \left(\hat{R}_{\omega,u} - \hat{\alpha}_u F(k)\right)\left(1 - \hat{R}_{\omega,u} + \hat{\alpha}_u F(k)\right).$$

In the majority of cases we examined, the least squares estimates
R_ω and α were better first approximations to $\hat{R}_\omega$ and $\hat{\alpha}$ in the sense
that they converged more rapidly with successive applications of (26)
and (27). Thus we present the least squares estimates of R_ω and α
for the sake of completeness; denoting these by R_ω^* and α^* we obtain,
with the aid of (13) and (14),

$$(29) \quad R_\infty^* = \frac{\left(\sum_{k=1}^{N} F^2(k)\right)\left(\sum_{k=1}^{N} s_k/n_k\right) - \left(\sum_{k=1}^{N} F(k)\right)\left(\sum_{k=1}^{N} F(k)s_k/n_k\right)}{N\sum_{k=1}^{N} F^2(k) - \left(\sum_{k=1}^{N} F(k)\right)^2},$$

and

$$(30) \quad \alpha^* = \frac{\left(\sum_{k=1}^{N} F(k)\right)\left(\sum_{k=1}^{N} s_k/n_k\right) - N\left(\sum_{k=1}^{N} F(k)s_k/n_k\right)}{N\sum_{k=1}^{N} F^2(k) - \left(\sum_{k=1}^{N} F(k)\right)^2}.$$

A third and more general approach of solving (22) and (23) is an application of Newton's method in two dimensions. For the sake of completeness we now present this method in general terms.

Suppose that $F(x, y)$ and $G(x, y)$ are, at least once, differentiable functions in the variables x and y. We then wish to solve, iteratively, the equations

$$(31) \quad F(x, y) = 0$$

and

$$(32) \quad G(x, y) = 0.$$

Suppose that x_0 and y_0 ar initial guesses to the solution. Then, the n-th iteration to the solution is given as

$$(33) \quad \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} - \begin{bmatrix} \dfrac{\partial F(x,\ y)}{\partial x} & \dfrac{\partial F(x,\ y)}{\partial y} \\ \dfrac{\partial G(x,\ y)}{\partial x} & \dfrac{\partial G(x,\ y)}{\partial y} \end{bmatrix}_{\substack{x=x_{n-1} \\ y=y_{n-1}}}^{-1} \begin{bmatrix} F(x_{n-1},\ y_{n-1}) \\ G(x_{n-1},\ y_{n-1}) \end{bmatrix}$$

in matrix notation. Individually, then,

$$(34) \quad x_n = x_{n-1} - \frac{\left[F(x,\ y)\ \dfrac{\partial G(x,\ y)}{\partial y} - \dfrac{\partial F(x,\ y)}{\partial y}\ G(x,\ y) \right]}{\left[\dfrac{\partial F(x,\ y)}{\partial x}\ \dfrac{\partial G(x,\ y)}{\partial y} - \dfrac{\partial F(x,\ y)}{\partial y}\ \dfrac{\partial G(x,\ y)}{\partial x} \right]}\Bigg|_{\substack{x=x_{n-1} \\ y=y_{n-1}}},$$

and

$$(35) \quad y_n = y_{n-1} - \frac{\left[G(x,\ y)\ \dfrac{\partial F(x,\ y)}{\partial x} - F(x,\ y)\ \dfrac{\partial G(x,\ y)}{\partial x} \right]}{\left[\dfrac{\partial F(x,\ y)}{\partial x}\ \dfrac{\partial G(x,\ y)}{\partial y} - \dfrac{\partial F(x,\ y)}{\partial y}\ \dfrac{\partial G(x,\ y)}{\partial x} \right]}\Bigg|_{\substack{x=x_{n-1} \\ y=y_{n-1}}}$$

The large sample 2 by 2 variance-covariance matrix for $\hat{R}_{\infty}$ and $\hat{\alpha}$ is

$$\sum = \begin{bmatrix} - E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_{\infty}^2} \right) & - E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial R_{\infty}\partial\alpha} \right) \\ & - E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial\alpha^2} \right) \end{bmatrix}^{-1} .$$

The second partial derivatives of $\log_e \mathcal{L}$ with respect to $\hat{R}_\infty$ and $\hat{\alpha}$ are given by (18), (19), and (20).

Since $\hat{R}_\infty$ and $\hat{\alpha}$ are approximately distributed normally with means R_∞ and α, respectively, and the variance-covariance matrix given previously, an approximate 100τ-percent lower confidence limit for R_k (the predicted reliability at the k-th stage of testing) is given by

(36) $\qquad \ell_k = \hat{R}_k - Z_{1-\tau} \sqrt{\text{Var } \hat{R}_k}$

$\qquad\qquad = \hat{R}_k - Z_{1-\tau} \sqrt{\text{Var } \hat{R}_\infty + F^2(k) \text{ Var } \hat{\alpha} - 2F(k) \text{ cov } (\hat{R}_\infty, \hat{\alpha})}$,

where $\text{Var } \hat{R}_\infty$, $\text{Var } \hat{\alpha}$ and $\text{cov } (\hat{R}_\infty, \hat{\alpha})$ are the elements from the matrix $\sum$, and $z_{1-\tau}$ is obtained from

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_{1-\tau}} e^{-z^2/2} \, dz = 1 - \tau.$$

Since theoretical values will not be available in practice, the maximum likelihood estimates $\hat{R}_\infty$ and $\hat{\alpha}$ are substituted in these equations to obtain numerical results, which appear in Sec. III.

As frequently happens with data analyses, (21) may yield a maximum likelihood estimate of R_∞ that is greater than unity. If this should occur, R_∞ is set equal to one and the function to be maximized is:

(37) $\qquad \mathcal{L} = \prod_{k=1}^{N} \binom{n_k}{s_k} \left[1 - \alpha F(k) \right]^{s_k} \left[\alpha F(k) \right]^{n_k - s_k}$,

$$\log_e \mathcal{L} = \text{const.} + \sum_{k=1}^{N} s_k \log_e \left[1 - \alpha F(k)\right] + \sum_{k=1}^{N} (n_k - s_k) \log_e \alpha,$$

(38)
$$\frac{\partial \log_e \mathcal{L}}{\partial \alpha} = -\sum_{k=1}^{N} \frac{s_k F(k)}{\left[1 - \alpha F(k)\right]} + \sum_{k=1}^{N} \frac{(n_k - s_k)}{\alpha}.$$

The maximum likelihood estimate is that α such that

(39)
$$-\sum_{k=1}^{N} \frac{s_k F(k)}{\left[1 - \hat{\alpha} F(k)\right]} + \sum_{k=1}^{N} \frac{(n_k - s_k)}{\hat{\alpha}} = 0.$$

To iterate on a solution for $\hat{\alpha}$, we let $\hat{\alpha}_n$ be the n-th iteration; then using Newton's method, we find

(40)
$$\hat{\alpha}_{n+1} = \hat{\alpha}_n + \frac{\left\{\dfrac{1}{\hat{\alpha}_n} \sum_{k=1}^{N} (n_k - s_k) - \sum_{k=1}^{N} \dfrac{s_k F(k)}{\left[1 - \hat{\alpha}_n F(k)\right]}\right\}}{\left\{\dfrac{1}{\hat{\alpha}_n^2} \sum_{k=1}^{N} (n_k - s_k) + \sum_{k=1}^{N} \dfrac{s_k F^2(k)}{\left[1 - \hat{\alpha}_n F(k)\right]^2}\right\}}.$$

MODEL 2

This model is utilized in Reliability Growth of U.S. Rockets (U) [9]; the model is not classified. In this case,

(41)
$$f(R_\infty, \alpha_1, \alpha_2, k) = 1 - \alpha_1 e^{-\alpha_2 k}.$$

That is, P_∞ is assumed to be unity in this case (which, depending upon the particular application, may or may not be a reasonable assumption). The likelihood function is then

$$(42) \qquad \mathcal{L} = \prod_{k=1}^{N} \binom{n_k}{s_k}\left[1 - \alpha_1 e^{-\alpha_2 k}\right]^{s_k}\left[\alpha_1 e^{-\alpha_2 k}\right]^{n_k-s_k}.$$

It is not difficult to show that $\log_e \mathcal{L}$ is a concave function in α_1 and α_2, insuring that the maximum is unique. We follow the approach in [9], assuming the maximum occurs in the region $0 < \alpha_1 < e^{\alpha_2}$. We now solve the equations

$$(43) \qquad \frac{\partial \log_e \mathcal{L}}{\partial \alpha_1} = \sum_{k=1}^{N} \frac{n_k\left(1 - s_k/n_k - \hat{\alpha}_1 e^{-\hat{\alpha}_2 k}\right)}{\hat{\alpha}_1\left(1 - \hat{\alpha}_1 e^{-\hat{\alpha}_2 k}\right)} = 0,$$

and

$$(44) \qquad \frac{\partial \log_e \mathcal{L}}{\partial \alpha_2} = \sum_{k=1}^{N} \frac{k n_k\left[\hat{\alpha}_1 e^{-\hat{\alpha}_2 k} - (1 - s_k/n_k)\right]}{\left(1 - \hat{\alpha}_1 e^{-\hat{\alpha}_2 k}\right)} = 0.$$

Solution in terms of $\hat{\alpha}_1$ and $\hat{\alpha}_2$ is accomplished by repeated use of equations (34) and (35).

The large sample 2 by 2 variance-covariance matrix for $\hat{\alpha}_1$ and $\hat{\alpha}_2$ is

$$\sum = \left[\begin{array}{cc} - E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1^2}\right) & - E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1 \partial \alpha_2}\right) \\ & - E\left(\dfrac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1^2}\right) \end{array} \right]^{-1} ,$$

where:

$$(45) \qquad \frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1^2} = \frac{1}{\alpha_1^2} \sum_{k=1}^{N} n_k \left[\frac{s_k/n_k \left(1 - 2\alpha_1 e^{-\alpha_2 k}\right)}{\left(1 - \alpha_1 e^{-\alpha_2 k}\right)^2} - 1 \right] ,$$

$$\frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1 \partial \alpha_2} = \sum_{k=1}^{N} \frac{k s_k e^{-\alpha_2 k}}{\left(1 - \alpha_1 e^{-\alpha_2 k}\right)^2} ,$$

and

$$(46) \qquad \frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_2^2} = -\alpha_1 \sum_{k=1}^{N} \frac{k^2 s_k e^{-\alpha_2 k}}{\left(1 - \alpha e^{-\alpha_2 k}\right)^2} .$$

Expanding R_k in a Taylor series in α_1 and α_2 up to and including terms of the first order, a special case of (15), we approximate the variance of R_k for each k by the following expression, which appears in [9]:

$$(47) \qquad \mathrm{Var}\ R_k \doteq e^{-2\alpha_2 k} \left[\mathrm{Var}\ \hat{\alpha}_1 + \alpha_1^2 k^2\ \mathrm{Var}\ \hat{\alpha}_2 + 2\alpha_1 k\ \mathrm{cov}\ (\hat{\alpha}_1,\ \hat{\alpha}_2) \right] .$$

While large-sample theory yields asymptotic normality of $\hat{R}_k$, simulation results indicate that for the number of trials which might reasonably be expected in a development program, the beta distribution provides a better (in many cases, nearly perfect) approximation. If it is assumed that the parameters of the beta distribution are p_k and q_k, by the method of moments, we find that

$$
(48) \qquad p_k = R_k \left[\frac{R_k(1 - R_k)}{\operatorname{Var} R_k} - 1 \right] ,
$$

and

$$
(49) \qquad q_k = (1 - R_k) \left[\frac{R_k(1 - R_k)}{\operatorname{Var} R_k} - 1 \right] .
$$

Thus a 100τ-percent lower confidence limit for R_k is given by $L_{k,\tau}$, where L_k is the solution to the equation

$$
(50) \qquad \int_0^{L_{k,\tau}} \frac{\Gamma(p_k + q_k)}{\Gamma(p_k)\Gamma(q_k)} x^{p_k - 1} (1 - x)^{q_k - 1} \, dx = 1 - \tau .
$$

This development is found in [9].

If the maximum likelihood estimate of α_1 is negative, β is set equal to zero and

$$
(51) \qquad R_k = (1 - \alpha_1),
$$

in which there is no reliability growth. The corresponding estimation problem is then not very interesting or enlightening and can be done without difficulty.

If the maximum likelihood estimate of α_1 causes R_k to become negative, we then put $\alpha_1 = 1$. Thus,

$$(52) \qquad R_k = 1 - e^{-\alpha_2 k},$$

and the function we maximize is

$$(53) \qquad \mathcal{L} = \prod_{k=1}^{n} \binom{n_k}{s_k} \left(1 - e^{-\alpha_2 k}\right)^{s_k} \left(e^{-\alpha_2 k}\right)^{n_k - s_k},$$

$$(54) \qquad \log_e \mathcal{L} = \text{const.} + \sum_{k=1}^{N} s_k \left(1 - e^{-\alpha_2 k}\right) - \sum_{k=1}^{N} \alpha_2 k \left(n_k - s_k\right).$$

Thus,

$$(55) \qquad \frac{\partial \log_e \mathcal{L}}{\partial \alpha_2} = \sum_{k=1}^{N} \frac{s_k e^{-\alpha_2 k}}{\left(1 - e^{-\alpha_2 k}\right)} - \sum_{k=1}^{N} k(n_k - s_k).$$

The maximum likelihood estimator $\hat{\alpha}_2$ is then that value of α_2 such that

$$(56) \qquad \sum_{k=1}^{N} \frac{s_k e^{-\hat{\alpha}_2 k}}{\left(1 - e^{-\hat{\alpha}_2 k}\right)} - \sum_{k=1}^{N} k(n_k - s_k) = 0,$$

or

$$\sum_{k=1}^{N} \frac{s_k}{\left(e^{\hat{\alpha}_2 k} - 1\right)} - \sum_{k=1}^{N} k(n_k - s_k) = 0. \tag{57}$$

Again using Newton's method to iterate on a solution for $\hat{\alpha}_2$, we have the following: if $\hat{\alpha}_{2,n-1}$ is the $(n-1)$st iteration for $\hat{\alpha}_2$, then

$$\hat{\alpha}_{2,n} = \hat{\alpha}_{2,n-1} + \left[\frac{\sum_{k=1}^{N} \frac{s_k}{\left(e^{\hat{\alpha}_{2,n-1}} - 1\right)} - \sum_{k=1}^{N} k(n_k - s_k)}{\sum_{k=1}^{N} \frac{k s_k e^{\hat{\alpha}_{2,n-1} k}}{\left(e^{\hat{\alpha}_{2,n-1}} - 1\right)^2}} \right]. \tag{58}$$

MODEL 3

This is one of the two models introduced here. We assume in this case that

$$f(R_\infty, \alpha, k) = R_\infty - \alpha^k, \tag{59}$$

for the region $0 < \alpha^k < R_\infty < 1$, for $k = 1, \ldots, N$; thus for the region $0 < \alpha < R_\infty < 1$.

From (8) and (9) we have

$$\frac{\partial \log_e f}{\partial R_\infty} = \sum_{k=1}^{N} \frac{s_k}{(R_\infty - \alpha^k)} - \sum_{k=1}^{N} \frac{(n_k - s_k)}{(1 - R_\infty + \alpha^k)} = 0, \tag{60}$$

$$\frac{\partial \log_e f}{\partial \alpha} = - \sum_{k=1}^{N} \frac{k s_k \alpha^{k-1}}{(R_\infty - \alpha^k)} + \sum_{k=1}^{N} \frac{k(n_k - s_k)\alpha^{k-1}}{(1 - R_\infty + \alpha^k)} = 0. \tag{61}$$

The problem that arises here and that is brought out more clearly when we apply this model to actual reliability predictions is that we have no guarantee (60) and (61) will have unique solutions thereby enhancing the difficulty in obtaining the maximum likelihood estimates. This is because $\log_e \mathcal{L}$ is not necessarily a concave function of R_∞ and α.

To demonstrate this, the second partial derivatives with respect to $\log_e \mathcal{L}$ are:

$$(62) \quad \frac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty^2} = -\left[\sum_{k=1}^{N} \frac{s_k}{(R_\infty - \alpha^k)^2} + \sum_{k=1}^{N} \frac{(n_k - s_k)}{(1 - R_\infty + \alpha^k)^2} \right],$$

$$(63) \quad \frac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty \partial \alpha} = \left[\sum_{k=1}^{N} \frac{k s_k \alpha^{k-1}}{(R_\infty - \alpha^k)^2} - \sum_{k=1}^{N} \frac{k s_k \alpha^{k-1}}{(1 - R_\infty + \alpha^k)^2} \right],$$

and

$$(64) \quad \frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha^2} = \sum_{k=1}^{N} k(k-1)\alpha^{k-2} \left[\frac{(n_k - s_k)}{(1 - R_\infty + \alpha^k)} - \frac{s_k}{(R_\infty - \alpha^k)} \right]$$

$$- \sum_{k=1}^{N} s_k \left[\frac{k\alpha^{k-1}}{(R_\infty - \alpha^k)} \right]^2 - \sum_{k=1}^{N} (n_k - s_k) \left[\frac{k\alpha^{k-1}}{(1 - R_\infty + \alpha^k)} \right]^2,$$

and it follows that sufficient conditions guaranteeing a unique maximum, by solving the equations $\frac{\partial \log_e \mathcal{L}}{\partial R_\infty} = 0$, and $\frac{\partial \log_e \mathcal{L}}{\partial \alpha} = 0$, are violated when

$$\frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha^2} > 0.$$

Reducing (64), we see that

$$\frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha^2} < 0$$

if and only if

(65) $$\sum_{k=1}^{N} \frac{k\alpha^{k-2} s_k}{(R_\infty - \alpha^k)^2} \left[(k-1)R_\infty + \alpha^k \right]$$

$$> \sum_{k=1}^{N} \frac{k\alpha^{k-2}(n_k - s_k)}{(1 - R_\infty + \alpha^k)^2} \left[(k-1)(1 - R_\infty) - \alpha^k \right].$$

We demonstrate this, by obtaining the second partial derivatives

$$\frac{\partial^2 \log_e \mathcal{L}}{\partial R_\infty^2}, \quad \frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha \partial R_\infty}, \quad \text{and} \quad \frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha^2}$$

(66) $$\log_e \mathcal{L} = \text{const.} + \sum_{k=1}^{N} s_k \left(\log_e \alpha_1 - \alpha_2/k \right)$$

$$+ \sum_{k=1}^{N} (n_k - s_k) \log_e \left(1 - \alpha_1 e^{-\alpha_2/k} \right),$$

(67) $$\frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1^2} = -\sum_{k=1}^{N} \left[s_k/\alpha_1^2 + \frac{(n_k - s_k)e^{-\alpha_2/k}}{\left(1 - \alpha_1 e^{-\alpha_2/k} \right)^2} \right],$$

$$(68) \quad \frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_1 \partial \alpha_2} = \sum_{k=1}^{N} \frac{(n_k - s_k)e^{-\alpha_2/k}}{k\left(1 - \alpha_1 e^{-\alpha_2/k}\right)^2} ,$$

and

$$(69) \quad \frac{\partial^2 \log_e \mathcal{L}}{\partial \alpha_2^2} = -\sum_{k=1}^{N} \frac{(n_k - s_k)\alpha_1 e^{-\alpha_2/k}}{k^2\left(1 - \alpha_1 e^{-\alpha_2/k}\right)^2} .$$

Analogous to the discussion of Model 3, we can state the following result. If α_1 and α_2 are so restricted that

$$(70) \quad \left\{ \sum_{k=1}^{N} \left[s_k/\alpha_1^2 + \frac{(n_k - s_k)e^{-2\alpha_2/k}}{\left(1 - \alpha_1 e^{-\alpha_2/k}\right)^2} \right] \right\} \cdot$$

$$\left[\sum_{k=1}^{N} \frac{(n_k - s_k)\alpha_1 e^{-\alpha_2/k}}{k^2\left(1 - \alpha_1 e^{-\alpha_2/k}\right)^2} \right] \le \left[\sum_{k=1}^{N} \frac{(n_k - s_k)e^{-\alpha_2/k}}{k\left(1 - \alpha_1 e^{-\alpha_1/k}\right)^2} \right]^2 ,$$

sufficient conditions guaranteeing a unique maximum by solving

$$\frac{\partial \log_e \mathcal{L}}{\partial \alpha_1} = \frac{\partial \log_e \mathcal{L}}{\partial \alpha_2} = 0$$

are violated.

CONFIDENCE INTERVALS FOR MODELS 3 AND 4

We can rewrite Model 3 in the alternative form

$$(71) \qquad R_k = R_\infty - e^{-\beta k},$$

where $\beta = -\log_e \alpha$. Thus, by the invariance principle of maximum likelihood estimators

$$\hat{R}_k = \hat{R}_\infty - e^{-\hat{\beta} k},$$

where $\hat{\beta} = -\log_e \hat{\alpha}$. By (15),

$$(72) \qquad \text{Var } \hat{R}_k \doteq \text{Var } \hat{R}_\infty + k^2 e^{-2\beta k} \text{ Var } \hat{\beta} - 2k e^{-\beta k} \text{ cov } (\hat{R}_\infty, \hat{\beta}).$$

Thus, an approximate 100τ-percent lower confidence limit for R_k (the predicted reliability at the k-th stage of testing) is given by

$$(73) \quad \ell_k = \hat{R}_k - Z_{1-\tau} \sqrt{\text{Var } \hat{R}_k}$$

$$= \hat{R}_k - Z_{1-\tau} \sqrt{\text{Var } \hat{R}_\infty + k^2 e^{-2\beta k} \text{ Var } \hat{\beta} - 2k e^{-\beta k} \text{ cov } (\hat{R}_\infty, \hat{\beta})},$$

where $\text{Var } \hat{R}_\infty$, $\text{cov } (\hat{R}_\infty, \hat{\beta})$ and $\text{Var } \hat{\beta}$ are obtained in the usual manner and $Z_{1-\tau}$ is the solution of the equation

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_{1-\tau}} e^{-z^2/2} \, dz = 1 - \tau.$$

The development of the lower confidence limit for Model 4 is completely analogous with a different value of $\text{Var } \hat{R}_k$. For Model 4

$$(74) \qquad \text{Var } \hat{r}_k \doteq e^{-2\alpha_2/k} \left[\text{Var } \hat{\alpha}_1 + \frac{\alpha_1^2}{k^2} \text{Var } \hat{\alpha}_2 - \frac{2\alpha_1}{k} \text{cov } (\alpha_1, \alpha_2) \right].$$

ALTERNATIVE BAYESIAN PROCEDURE FOR
OBTAINING LOWER CONFIDENCE INTERVALS

The test data are divided into N stages by some predetermined criterion such as design change. Since flight test data are of the go, no-go variety, the probability that exactly x_1 successes are recorded in n_1 trials at Stage 1 is

$$(75) \qquad f_{11}(x_1|R_1) = \binom{n_1}{x_1} R_1^{x_1} (1 - R_1)^{n_1 - x_1}, \qquad x_1 = 0, 1, \ldots, n_1.$$

We shall assume that R_1 has a beta density with parameters a_1 and b_1 as its prior probability density function (pdf). That is,

$$(76) \qquad f_{12}(R_1) = \begin{cases} \dfrac{\Gamma(a_1 + b_1)}{\Gamma(a_1)\Gamma(b_1)} R_1^{a_1-1} (1 - R_1)^{b_1-1}, & 0 \leq R_1 \leq 1. \\ \\ 0, & \text{elsewhere.} \end{cases}$$

If we have no previous information regarding the system's reliability at Stage 1, we must choose the parameters a_1 and b_1 subjectively. We shall use the method Fox [4] describes. First of all, we let $\hat{R}_1$ be our subjective estimate of R_1. If we agree that our subjective estimate should be our most likely estimate we then set

$$(77) \qquad R_1 = \frac{\hat{a}_1 - 1}{\hat{a}_1 + \hat{b}_1 - 2},$$

assuming that (77) is not the uniform prior density on (0, 1). (If (77) is taken as the uniform prior density the modification to be made is discussed at the end of the subsection.) We then ask what the odds are that the true value of R_1 will lie in the interval $(\hat{R}_1 - k\hat{R}_1, \hat{R}_1 + k\hat{R}_1)$, where k is predetermined? If we set these odds at x to y, we can express this mathematically as

$$(78) \qquad \int_{\hat{R}_1 - k\hat{R}_1}^{\hat{R}_1 + k\hat{R}_1} f_{12}(R_1) \ dR_1 = v,$$

where $v = \dfrac{x}{x + y}$. This is Fox's equation [4, Eq. (1.6), p. 3]. From (77) and (78), with the aid of the tables in [4], we can now determine $\hat{a}_1$ and $\hat{b}_1$.

To obtain a Bayesian lower confidence interval for R_1 after we have completed Stage 1, we note that the posterior pdf of R_1, given x_1, is defined by

$$(79) \quad f_{13}(R_1|x_1) = \frac{\Gamma(\hat{a}_1 + \hat{b}_1 + n_1)}{\Gamma(\hat{a}_1 + x_1)\Gamma(\hat{b}_1 + (n_1 - x_1))} R_1^{\hat{a}_1 + x_1 - 1} (1 - R_1)^{\hat{b}_1 + (n_1 - x_1) - 1}.$$

Thus, a $(1 - \alpha_1)$ - 100-percent level Bayesian lower confidence interval for R_1 is given by

$$(80) \qquad \int_{\xi_1}^{1} f_{13}(R_1|x_1) \ dR_1 = \alpha_1,$$

where ξ_1 is so chosen that (80) holds.

Following the method of derivation for Stage 1, we can now obtain a $(1 - \alpha_2) \bullet 100$-percent level Bayesian lower confidence interval for R_2 in an analogous manner. We use the data in Stage 1, however, to require that the most likely estimate for R_2 is given by

$$(81) \qquad R_2 = \frac{\hat{a}_2 - 1}{\hat{a}_2 + \hat{b}_2 - 2} \geq \frac{\hat{a}_1 + x_1 - 1}{\hat{a}_1 + \hat{b}_1 + n_1 - 2} .$$

Thus, by this method we obtain successive Bayesian estimates of reliability growth at each stage and their associated lower confidence bounds. There appears to be no way in which we can use the Bayesian approach to obtain a lower confidence bound at the $(N + 1)$th stage (i.e., for the predicted reliability at Stage $N + 1$ of our testing program).

The procedure described above needs slight modification if we assume a uniform density for R_1. (In this case, we assume total ignorance of the system's reliability prior to Stage 1. That is, any one value of R_1 is as likely to occur as any other value prior to starting the test program.) In this instance we know that $a_1 = b_1 = 1$, which in turn yields

$$(82) \qquad f_{13}(R_1|x_1) = \frac{\Gamma(n_1 + 2)}{\Gamma(x_1 + 1)\Gamma(n_1 - x_1 + 1)} R_1^{x_1}(1 - R_1)^{n_1 - x_1}.$$

We then proceed as before to obtain confidence intervals at Stage 1 and the succeeding stages for which data are available.

PRECEDING
PAGE BLANK

REFERENCES

1. Barlow, R. E., and E. M. Scheuer, _Reliability Growth During a Development Testing Program_, The RAND Corporation, RM-4317-1-NASA, October 1965.

2. Bresenham, J. E., "Reliability Growth Models," Technical Report 74, Department of Statistics, Stanford University, August 1964.

3. Cramér, H., _Mathematical Methods of Statistics_, 6th ed., Princeton University Press, Princeton, New Jersey, 1964.

4. Fox, B. L., _A Bayesian Approach to Reliability Assessment_, The RAND Corporation, RM-5084-NASA, August 1966.

5. Kendall, J. G., _The Advanced Theory of Statistics_, McGraw-Hill, New York, 1950.

6. Lloyd, D. K., and M. Lipow, _Reliability: Management Methods, and Mathematics_, Prentice-Hall Space Technology Series, Englewood Cliffs, New Jersey, 1962.

7. Weiss, H. K., "Estimation of Reliability Growth in a Complex System with a Poisson-Type Failure," _Operations Research_, Vol. 4, 1956.

8. Wolman, W., "Problems in System Reliability Analysis," in M. Zelen (ed.), _Statistical Theory of Reliability_, University of Wisconsin Press, Madison, 1963, pp. 149-160 (see also comments by Captain W. J. Corcoran, USN, p. 164).

9. Space Technology Laboratories, Inc., _Reliability Growth of U.S. Rockets_ (U), Vol. II, prepared for NASA/MSFC under Contract No. NAS 8-11037, May 18, 1964 (SECRET).

10. Howard, W. J., _Summary of Reliability Data for Developmental Missiles_, The RAND Corporation, RM-786-PR, April 1952 (FOR OFFICIAL USE ONLY).

11. -----, _Reliability of Nike and Terrier_, The RAND Corporation, RM-1247, May 1954 (FOR OFFICIAL USE ONLY).

DOCUMENT CONTROL DATA

1. ORIGINATING ACTIVITY	2a. REPORT SECURITY CLASSIFICATION
THE RAND CORPORATION	UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE   RELIABILITY ASSESSMENT IN THE PRESENCE OF RELIABILITY GROWTH

4. AUTHOR(S) (Last name, first name, initial)

Gross, A. J. and M. Kamins

5. REPORT DATE	6a. TOTAL No. OF PAGES	6b. No. OF REFS.
September 1967	66	9

7. CONTRACT OR GRANT No.	8. ORIGINATOR'S REPORT No.
F44620-67-C-0045	RM-5346-PR

9a. AVAILABILITY/LIMITATION NOTICES	9b. SPONSORING AGENCY
DDC-1	United States Air Force Project RAND

10. ABSTRACT

A methodology for estimating current and future reliability of complex weapon systems that show reliability growth during their development and early operational phases. The study proposes four reliability growth models or patterns that can be fitted to actual data experience to determine the quantitative characteristics of the growth within relatively well-defined tolerances. This is achieved by defining appropriate parametric models and subsequently using maximum likelihood procedures to obtain estimates of the parameters. Comparison of the models shows that under the conditions set forth in this study, three are generally superior in their predictive and assessment characteristics to representative non-parametric methods and to an applicable Bayesian procedure.

11. KEY WORDS

Weapon systems
Reliability
Numerical methods and processes
Testing
Research and development